

The star formation rate distribution function of the local Universe

M. S. Bothwell,^{1*} R. C. Kenicutt,¹ B. D. Johnson,¹ Y. Wu,² J. C. Lee,³ D. Dale,⁴
C. Engelbracht,⁵ D. Calzetti⁶ and E. Skillman⁷

¹*Institute of Astronomy, University of Cambridge, Cambridge CB3 0HA*

²*Infrared Processing and Analysis Center, California Institute of Technology, 1200 E. California Blvd, Pasadena, CA 91125, USA*

³*Carnegie Observatories of Washington, Pasadena, CA 91101, USA*

⁴*Department of Physics & Astronomy, University of Wyoming, Laramie, WY 82071, USA*

⁵*Steward Observatory, University of Arizona, Tucson, AZ 85721, USA*

⁶*Department of Astronomy, University of Massachusetts, Amherst, MA 01003, USA*

⁷*Department of Astronomy, School of Physics and Astronomy, University of Minnesota, 116 Church Street S.E., Minneapolis, MN 55455, USA*

Accepted 2011 March 29. Received 2011 March 28; in original form 2010 November 9

ABSTRACT

We present total infrared (IR) and ultraviolet (UV) luminosity functions derived from large representative samples of galaxies at $z \sim 0$, selected at IR and UV wavelengths from the Imperial *IRAS* Faint Source Catalogue redshift data base (IIFSCz) catalogue, and the *GALEX* All-Sky Imaging Survey (AIS), respectively. We augment these with deep *Spitzer* and *GALEX* imaging of galaxies in the 11 Mpc Local Volume Legacy (LVL) Survey, allowing us to extend these luminosity functions to lower luminosities ($\sim 10^6 L_\odot$), and providing good constraints on the slope of the luminosity function at the extreme faint end for the first time. Using conventional star formation prescriptions, we generate from our data the star formation rate (SFR) distribution function for the local Universe. We find that it has a Schechter form, the faint-end slope has a constant value (to the limits of our data) of $\alpha = -1.51 \pm 0.08$ and the ‘characteristic’ SFR ψ^* is $9.2 M_\odot \text{ yr}^{-1}$. We also show the distribution function of the SFR volume density; we then use this to calculate a value for the total SFR volume density at $z \sim 0$ of $0.025 \pm 0.0016 M_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}$, of which ~ 20 per cent is occurring in starbursts. Decomposing the total star formation by infrared luminosity, it can be seen that 9 ± 1 per cent is due to LIRGs, and 0.7 ± 0.2 per cent is occurring in ULIRGs. By comparing UV and IR emission for galaxies in our sample, we also calculate the fraction of star formation occurring in dust-obscured environments, and examine the distribution of dusty star formation: we find a very shallow slope at the highly extinguished end, which may be attributable to line-of-sight orientation effects as well as conventional internal extinction.

Key words: galaxies: evolution – galaxies: luminosity function, mass function – galaxies: star formation – cosmology: observations.

1 INTRODUCTION

Measuring the distribution of star formation rates (SFRs) at $z = 0$ provides a fundamental observational reference point for galaxy evolution models, being a directly testable result of the hierarchical galaxy formation process. As such, obtaining good constraints on the local SFR distribution function is an integral part of calibrating and constraining cosmological models concerning the growth and buildup of structure in the Universe.

In an ideal world, it would be possible to directly measure the bolometric output of *all* galaxies in some large volume-limited sample, thus measuring – without biases – the true underlying distribu-

tion of star formation and luminosity. In reality, however, there are two major constraints to be overcome in order to approach the bias-free distribution functions. First, for large samples the bolometric output and SFR must be extrapolated from one or more luminosity components, typically $H\alpha$, ultraviolet (UV) or infrared (IR), which necessarily involves some uncertainty (particularly for extrapolations based on a single continuum wavelength; see e.g. Calzetti et al. 2000). Secondly, the sample’s selection function plays a large part in determining the behaviour of the derived distribution, with samples selected in the UV and IR probing very different populations with very different underlying star formation behaviour and dust content (e.g. Buat et al. 2005; Iglesias-Páramo et al. 2006). The behaviour of dust is also of great interest, as extinction, correlating strongly with star formation, plays a major role in the derived star formation properties of a sample.

*E-mail: bothwell@ast.cam.ac.uk

Measuring the luminosity function (LF) and the SFR distribution function has been an important goal for many years, with much work being dedicated towards successfully estimating the local LF (Takeuchi, Yoshikawa & Ishii 2003; Schiminovich et al. 2005; Wyder et al. 2005; Wang & Rowan-Robinson 2010), the SFR distribution (e.g. Barkana 2002; Martin et al. 2005; Martin et al. 2007) and the closely related total SFR volume density (Gallego et al. 1995; Tresse & Maddox 1998; Hopkins 2004). These studies, often based on the large homogenous *IRAS* and *GALEX* survey data sets, have provided good constraints at the bright, highly star-forming end of the distribution. However, the behaviour of the faint-end slope is still somewhat unexplored, with most studies suffering significant incompleteness at low luminosities ($\lesssim 10^{8-9} L_{\odot}$) and SFRs ($\lesssim 10^{-1} M_{\odot} \text{ yr}^{-1}$). These less well-explored populations at the faint end, being numerically dominant, provide a vital insight into the hierarchical galaxy formation process.

It is a well-established result that the low-mass slope of the predicted dark matter (DM) mass function is steeper than the observed faint-end slope of the LF: this discrepancy is often framed as the well-known ‘missing satellite problem’ (Klypin et al. 1999; Moore et al. 1999). The faint-end slope of the LF is a way of quantifying the number of faint galaxies, and while models adopting a Press–Schechter formalism predict a steep faint-end slope of $\alpha \sim -1.8$ (Press & Schechter 1974; Boylan-Kolchin et al. 2009), observational results suggest a significantly shallower value (Marinoni et al. 1999; Liske et al. 2003).

However, the form of the observed LF is several steps removed from the DM mass function, being the end result of many non-linear baryonic processes. A multitude of physical processes dictate how DM haloes collect and retain their gas, setting the form of the H I mass function which, through star formation, produces the LF. Measuring and comparing the distribution functions pertaining to these baryonic processes (i.e. the hydrogen mass function, the SFR function and the LF) has the potential to shed light on the possible reasons that baryons are underabundant in low-mass DM haloes, causing the shallow slope and ‘missing satellites’.

To attempt to obtain good constraints on the extreme faint end, we use data from the Local Volume Legacy (LVL) survey, a volume-limited data set of galaxies within 11 Mpc with deep multi-band *Spitzer*, optical and *GALEX* imaging. Being volume limited, this survey provides an excellent and unbiased data base from which to calculate IR and UV luminosities and SFRs, which can be combined with the same parameters from the larger data sets. For the first time, therefore, we are able to augment our large survey data sets with deep imaging of this statistically complete sample of local galaxies, and extend the local luminosity and SFR distributions functions 1–2 orders of magnitude deeper.

In this paper, we present statistically complete SFR distribution functions and luminosity functions which extend 1–2 orders of magnitude deeper than previous studies. The data provide the first ever SFR distribution study complete at both the bright and extreme faint ends, and give an unbiased view of the distribution of star formation and structure at $z \sim 0$. We use our samples to accurately constrain the true value – and breakdown – of the SFR volume density in the $z = 0$ Universe. In order to investigate the role of extinction (which correlates strongly with SFR), we also use IR and UV photometry to analyse the behaviour of dust-obscured star formation (measured via the ratio $L_{\text{IR}}/L_{\text{FUV}}$), and present ‘extinction distribution functions’, analogous to the LF.

Throughout this paper we assume a concordance Λ cold dark matter (Λ CDM) cosmology with parameters (H_0 , Ω_m , Ω_{Λ}) = (72 km s^{−1} Mpc^{−1}, 0.27, 0.73). The Local Volume sample, for which

peculiar motions overwhelm the Hubble flow, has distances based on a number of direct distance estimators wherever possible (see Section 2.2.2).

2 METHODOLOGY

2.1 Sample selection

As mentioned in Section 1 above, for any observationally defined sample the nature of the selection function will go a long way to determining the physical properties of the sample. Any non-volume-limited sample used to calculate an LF must be strictly flux limited at some wavelength, so that when weightings are applied (e.g. volume weightings, as per the $1/V_{\text{max}}$ method) the resulting function, $\Phi(L)$, becomes statistically representative of the underlying population as a whole (an inconsistent flux limit would lead to bias through inconsistent volume corrections).

The wavelength at which this limit is imposed will to a large extent determine the properties of the sample. An IR-selected sample (selected at 60 μm for example) which is then crossed with a UV data base will contain more dusty, obscured systems than an initially FUV-selected sample which is subsequently cross matched with an IR data base. Buat et al. (2007) demonstrated that the bolometric LFs for IR and UV-selected samples diverged for luminous systems ($L_{\text{bol}} > 5 \times 10^{10} L_{\odot}$), with a significant proportion of the massive, luminous systems detected strongly at 60 μm being non-detected in the FUV due to their extreme dust attenuation. At the low-luminosity end, samples suffer from the opposite bias; late-type, metal-poor, star-forming galaxies may lack sufficient dust to re-process the UV photons into IR emission, and as a result can be weak- or non-detections at 60 μm despite having measurable levels of ongoing star formation (Dale et al. 2009; Calzetti et al. 2010).

Clearly, defining a sample in terms of selection at a single wavelength and extrapolating global properties will lead to an inaccurate view of the behaviour of star formation and extinction in the overall population. In this paper, we therefore use separate data sets selected at 60 μm and *GALEX* FUV – hereafter, these will be referred to as the ‘IR-selected’ and ‘UV-selected’ samples, respectively. These dual samples can then be combined, and the resultant data set encapsulates the full range of star-forming behaviour, ensuring that both obscured and unobscured star formation are represented.

2.1.1 IR-selected sample

For our IR-selected sample, we make use of the Imperial *IRAS* Faint Source Catalogue redshift data base (IIFSCz) compiled from the *IRAS* Faint Source Catalogue (FSC) by Wang & Rowan-Robinson (2009). This is a large (61 per cent of the sky) data base consisting of 60,303 galaxies selected at 60 μm . For our analysis, we require that all galaxies have redshifts, and we include both spectroscopic and photometric redshifts in our compilation (Wang & Rowan-Robinson 2009 discuss in detail the accuracy of their photometric redshifts, derived using neural network fitting, concluding that they are accurate to 1.5–2.3 per cent). A total of 44 622 galaxies in the IIFSCz have a secure redshift identification. We then cross matched this subsample against the *GALEX* All-Sky Imaging Survey (AIS), resulting in a total of 25 768 galaxies observed with both IR and UV: this comprises our IR-selected sample. Naturally, some of the galaxies observed with *GALEX* were not detected due to a low UV flux. We distinguished between *GALEX* non-detections (i.e. upper limits) and galaxies never observed with *GALEX* by cross matching with a 0.6 search radius, equivalent to the size of a *GALEX* tile.

To ensure that redshift evolution effects played no significant part in the luminosity distribution, we applied a $z < 0.1$ cut to our sample (which will slightly reduce the number density of the most luminous systems, which are under-represented in the local Universe). We also applied a $60\text{ }\mu\text{m}$ flux limit of 0.36 Jy , which was then used to define the value of V_{max} for each galaxy in the IR sample – this limit corresponds to the 90 per cent completeness limit of the *IRAS* FSC, and is higher than the formal $60\text{ }\mu\text{m}$ flux limit of the FSC of 0.2 Jy . We also applied a minimum redshift cut of $z = 0.005$ ($\sim 20\text{ Mpc}$) to remove any galaxies for which peculiar motion would overwhelm the Hubble flow, leading to highly uncertain distances based upon recession velocity alone.

This last cut in effect removes IR-faint ($L_{\text{IR}} \leq 10^9 L_{\odot}$) galaxies from the sample; the volume-limited LVL sample samples this region of luminosity space well however, so no information is lost overall. This near field cut also has the effect of eliminating much of the *GALEX* ‘shredding’ problem whereby nearby extended sources are resolved as multiple objects, resulting in artificially lowered UV fluxes (see e.g. Salim et al. 2007). By comparing the UV fluxes of our IR-selected sample with those of Buat et al. (2007), who used their own photometric extraction technique, we estimate that at most ~ 5 per cent of our galaxies suffer from photometric extraction issues (such as shredding). Constructing LFs consisting solely of the subset of galaxies common to both samples (302 members), we find that the UV LF produced by our fluxes is essentially identical to that built from independently obtained UV fluxes. As such, we proceed with the assumption that our UV fluxes are robust.

The final IR sample, after applying all the above cuts, consists of 10 252 galaxies.

To confirm that working with this subsample of the IIFSCz does not bias our results, we constructed the (monochromatic) $60\text{ }\mu\text{m}$ LF for our subsample, and compared them to the equivalent LFs of Wang & Rowan-Robinson (2010) which were constructed using the complete sample; the LF derived from our subsample does not differ significantly from those derived from the parent sample (see Fig. 1 for the LF of our original, ‘unadulterated’ IR sample). The main difference is a sharper cut-off at the most luminous end of the

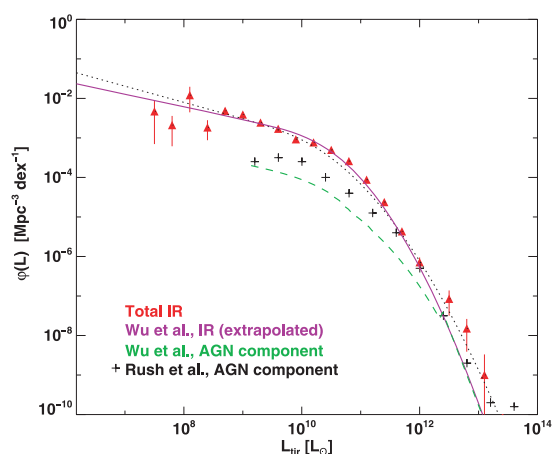


Figure 1. IR LF for our original IR-selected sample (red triangles), plotted with an extrapolated IR LF from Wu et al. (purple line). The AGN contribution to Wu et al.’s LF is shown as a green line. Plotted as black crosses is a previous estimate of the AGN contribution to the IR LF, from Rush et al. (1993). For reference, the canonical *IRAS* IR LF (Takeuchi et al. 2003) is plotted as a dotted black line. Note that our original IR-selected sample is shown here before any cuts, and is therefore identical to the LF of Wang & Rowan-Robinson (2010).

LF for our sample, which is attributable to the high- z cut – the parent sample, having no such cut, includes many ULIRGs and HyLIRGs (systems with $L_{\text{IR}} > 10^{12} L_{\odot}$ and $10^{13} L_{\odot}$, respectively), which are very rare in the $z \sim 0$ Universe.

2.1.2 UV-selected sample

The UV-selected sample was taken from the paper by Buat et al. (2007), who assembled a UV-selected sample of galaxies to assess the relative contributions of obscured and unobscured star formation in the local Universe. To briefly summarize, their sample was selected by applying a UV cut of $F_{\text{UV}} = 17.5\text{ mag}$ to the *GALEX* AIS catalogue, and cross-matching with the *IRAS* PSCz in areas uncontaminated by foreground (i.e. galactic) cirrus emission. The resulting effective area is 2210 deg^2 – while smaller than the large area covered by the IIFSCz (which, at 61 per cent of the sky, covers over an order of magnitude more sky area), this is still substantial and will avoid the clustering biases and sensitivity to small-scale inhomogeneities that are the weakness of ‘pencil beam’-type surveys. Distances for the UV sample galaxies were obtained from NASA/IPAC Extragalactic Data base (NED).

The same distance cuts as the IR-selected sample were applied ($0.005 < z < 0.1$). The low- and high- z cut-offs remove eight and four galaxies, respectively, from the parent sample of 606 – our UV-selected sample therefore consists of 595 galaxies.

2.1.3 The Local Volume sample

To augment our IR- and UV-selected samples in the low-luminosity regime, we use data on a complete sample of nearby galaxies collected by the *GALEX* 11 HUGS (11 Mpc $\text{H}\alpha$ UV Galaxy Survey) and *Spitzer* LVL programs. The sample is dominated by dwarf galaxies, and is thus ideal for studying the nature of systems with low SFRs, low metallicities and low dust contents. UV and mid- to far-IR flux catalogues are published in Lee et al. (2011) and Dale et al. (2010), respectively. Details on the sample selection, observations, photometry are provided in those papers, and in Kennicutt et al. (2008), who describe the overall parent 11 Mpc sample and $\text{H}\alpha$ imaging survey. A brief summary of the data set is given here.

The total parent Local Volume sample contains 436 objects. Galaxies are compiled from existing catalogues (as described in Kennicutt et al. 2008), and the selection is divided into two components. The primary component of the sample aims to be as complete as possible in its inclusion of known nearby star-forming galaxies within given limits. It consists of spirals and irregulars brighter than $B = 15\text{ mag}$ within 11 Mpc that avoid the Galactic plane ($|b| > 20^\circ$). These bounds represent the ranges within which the original surveys that provided the bulk of our knowledge on the Local Volume galaxy population have been shown to be relatively complete, while still spanning a large enough volume to probe a representative cross-section of star formation properties. The secondary component of the sample consists of galaxies that are within 11 Mpc and have available $\text{H}\alpha$ flux measurements, but fall outside one of the limits on brightness, Galactic latitude or morphological type. It is a composite of targets that either were observed by Kennicutt et al. (2008) as telescope time allowed or had $\text{H}\alpha$ fluxes published in the literature. Subsequent statistical tests, as functions of B -band apparent magnitudes and HI fluxes (compiled from the literature), show that the subset of galaxies with $|b| > 20^\circ$ is relatively complete to $M_B \lesssim -15$ and $M_{\text{H}\text{I}} > 2 \times 10^8 M_{\odot}$ at the edge of the 11 Mpc volume (Lee et al. 2009).

Subsequent *GALEX* UV imaging primarily targeted the $|b| > 30^\circ$, $B < 15.5$ subset of the sample. The more restrictive latitude limit was imposed to avoid excessive Galactic extinction and fields with bright foreground stars and/or high background levels for which observations would be prohibited due to *GALEX*'s brightness safety restrictions. Deep, single orbit (~ 1500 s) imaging was obtained for each galaxy, following the strategy of the *GALEX* Nearby Galaxy Survey (Gil de Paz et al. 2007). *GALEX* observations for a significant fraction of the remaining galaxies beyond these limits were also taken by other programs. Overall, *GALEX* data are available for ~ 90 per cent of the 436 galaxy sample.

Finally, *Spitzer* Infrared Array Camera (IRAC) mid-infrared and MIPS far-infrared imaging was also obtained for the $|b| > 30^\circ$, $B < 15.5$ subset of the sample through the LVL program. This subsample with both UV and IR coverage contains 257 galaxies and is used in the following analysis.

The resulting data set provides an unprecedented multi-wavelength view of star formation in the nearby Universe. In particular, the multi-band *Spitzer* IR observations allow for bolometric luminosities to be calculated (using the algorithms provided by Dale & Helou 2002) in a manner consistent with the *IRAS* fluxes available for our larger data sets.

2.2 Constructing the luminosity functions

2.2.1 $1/V_{\max}$ method

The ‘classic’ way of constructing an LF is based on the estimator

$$\Phi(L)dL = \sum_i \frac{dL}{V_{\max}(L_i, S_v^{\text{lim}})},$$

where V_{\max} is the volume enclosed by the maximum distance at which galaxy i would be observable, given the flux limits of the survey (S_v^{lim} – defined by the IR and UV, respectively, for the IR- and UV-selected samples, and the B band for the LVL sample) and the galaxy's luminosity – this ‘maximum volume’ is then used to weight each galaxies' contribution to the final function $\Phi(L)$. This inverse volume weighting is designed to counteract the Malmquist-type bias which would be encountered by a pure number counting exercise: the volume probed by a flux-limited survey varies as a function of luminosity, and as such faint galaxies (which are only seen nearby) are under-represented; conversely, bright galaxies seen out to large distances are overrepresented. Weighting by $1/V_{\max}$ eliminates this effect, and reconstructs the true underlying luminosity distribution. This particular form of the weighting (as opposed to first binning into luminosity bins, then applying a mean V_{\max} to all galaxies within the bin) reduces error resulting from binning of the data – in essence, each galaxy is assigned its own bin of width dL .

The $1/V_{\max}$ method has the great advantage of being a non-parametric estimator, as it does not assume any form of $\Phi(L)$. However, it does suffer from two weaknesses: first, binning in luminosity space is sometimes required, which necessarily involves some loss of information, and the resulting form of $\Phi(L)$ can be somewhat sensitive to the particular binning used. Secondly, and more seriously, it is highly sensitive to local density enhancements (see Efstathiou, Ellis & Peterson 1988 for a detailed discussion). The Local Volume represents a significant overdensity compared to the cosmic mean, which will manifest in a flux-limited survey as an enhancement at the faint end of the LF. Karachentsev et al. (2004) compared the B -band luminosity density in the local 8 Mpc and find that it is 1.7–2.0 times the global luminosity density (as derived

from both the Sloan Digital Sky Survey and the Millennium Galaxy Catalogue). We therefore correct the local density downwards by a factor of 1.85 ± 0.15 , and incorporate this uncertainty in the error on the derived faint-end slope. This is consistent with the values found by Lee et al. (2009b, 2011) by comparing the LVL and the field H mass function (the Local Volume is overdense by a factor of 1.4), the B -band LF (a factor of 2.3) and the UV LF (a factor of 2).

2.2.2 Maximum likelihood method

The weakness to density fluctuations can be overcome by adopting a parametric estimator, which assumes the form of the LF is universal (with the precise shape being determined by some free parameters), which allows the density to be factored out. One such statistic is the ‘maximum likelihood’ method, as described by, for example, Efstathiou et al. (1988) and Yahil et al. (1991).

We take a Schechter (1976) function to be the assumed form of the LF:

$$\Phi(L)dL = \Phi^* \left(\frac{L}{L^*} \right)^\alpha e^{-(L/L^*)} d \left(\frac{L}{L^*} \right).$$

When fitting this to data using a maximum likelihood method, the free parameters are obtained by maximizing the value of Δ with respect to α and L^* ,

$$\Delta = \sum_i \log F_i,$$

where

$$F_i = \frac{\Phi_i}{\Psi} = \frac{(L_i/L^*)^\alpha e^{-(L_i/L^*)}}{\Gamma(\alpha + 1, L_{\text{lim}}/L^*)},$$

and where $\Psi = \int \Phi(L)dL$ is the cumulative LF (equal to Γ , the normal incomplete gamma function for a Schechter LF), and L_{lim} is the minimum luminosity at which the sample is complete. Essentially, a grid of Schechter functions with varying parameters is explored, and each is assigned a ‘likelihood’ using the data. Our adopted function is taken to be the function which maximizes the ‘likelihood’.

While this method has the aforementioned advantage that it is insensitive to local density fluctuations, it has the disadvantage of being similarly insensitive to the absolute normalization of the LF (note that Φ^* cancels out in the expression for F_i). This can be recovered using number counting, as the integral under the derived best-fitting function is just the number of galaxies:

$$N(>L_{\text{lim}}) = \int_{L_{\text{lim}}}^{\infty} \Phi(L')dL = \Phi^* \Gamma(\alpha + 1, L_{\text{lim}}/L^*).$$

The uncertainty on the parameters of the LF – derived using either method – can be difficult to estimate. A major source of potential error, particularly at the faint end, is uncertainty on the distance. Kennicutt et al. (2008) discuss the distance estimates for the Local Volume galaxies in detail: direct distance estimators are used wherever possible (Cepheids, red giant branch, Tully–Fisher relation, supernovae), and when a distance from one of these methods is not available distances are estimated from the Hubble velocity, corrected for a Local Group dipole effect as derived by Karachentsev & Makarov (1996). The uncertainties on the distance (both random and systematic to the flow model) are typically 7–15 per cent.

The effect of this distance uncertainty (and the effect of the photometric flux uncertainty) will be to shift galaxies into neighbouring bins of luminosity (or SFR). As a result, the errors on the various

LF parameters are highly correlated, and have a complex dependence on the errors in the raw data. We deal with these errors using a Monte Carlo resampling method by creating a large number of realizations of the LF drawn – with replacement – from the parent sample. Parameters with well-defined errors (i.e. flux) were allowed to vary randomly according to a Gaussian likelihood function defined by the 1σ parameter error, and the standard deviation in the resultant LFs in each bin was taken to be an estimate of the error.

3 ANALYSIS

3.1 Removing the AGN contribution

Obscured active galactic nuclei (AGN) are a potential source of contamination for our sample, if we want to interpret the emission in terms of pure star formation. Unified models of AGN assume a dusty torus – which emits strongly in the IR – surrounding a central engine; this bright IR emission can be falsely interpreted as originating from dust heated by star-forming regions, which leads to an overestimation of the SFR of the galaxy. If we are to accurately estimate the SFR density of the volume containing our sample, it is important to remove such sources of possible contamination.

AGN number amongst the most extreme objects in the Universe, and, when active, have prodigious bolometric outputs. The effect of AGN subtraction will, therefore, be to reduce the very upper end of the LF, while leaving the behaviour at $L < 10^{10-11} L_{\odot}$ relatively unchanged (see e.g. Yuan, Kewley & Sanders 2010; Hopkins et al. 2010). At the upper end, star formation occurs almost entirely in highly extinguished, dusty environments (Calzetti et al. 2010 and references therein), where IR is by far the best SFR tracer: we therefore consider the effect of AGN contamination on our IR-selected sample only, as the UV-selected and LVL data sets will have a negligible AGN contribution (Salim et al. 2007).

It is difficult to perfectly remove all AGN on a galaxy-by-galaxy basis without detailed spectroscopy being available for the entire sample, using which the AGN component can be separated from the star-forming component (see e.g. Fu et al. 2010). In the absence of spectral data for our entire sample, a better method is to remove the AGN component statistically from the sample as a whole. This requires the knowledge of the global behaviour of the AGN population, in the form of an AGN LF which can then be subtracted from our original ‘total’ LF, leaving only the star-forming component.

To this end, we use the data in the work by Wu et al. (2011), a spectroscopic study of a subset of the 24 μm selected 5 mJy Unbiased Spitzer Extragalactic Survey (5MUSES) sample (Wu et al. 2011). This subset focused on galaxies with $z < 0.3$ from 5MUSES, with $\langle z \rangle \sim 0.12$. The 226 objects in this subsample were analysed with mid- and far-IR photometry (allowing a full characterization of their IR spectral energy distribution (SED)), and mid-IR spectroscopy – allowing each source’s luminosity to be decomposed into AGN and star-forming components. For the purpose of this analysis, we will examine the IR LF of their sample and use the relations given in their work to derive the contribution to the IR LF from AGN. We then take this fractional AGN contribution (as a function of IR luminosity) and subtract it from our data, leaving us with IR data for our galaxies which can be interpreted in terms of pure star formation.

Fig. 1 shows the LF for our IR-selected sample. Also plotted is the total IR LF for the galaxies reported in Wu et al. (2011), derived by taking their $15 \mu\text{m}$ LF and transforming into an IR LF using the average luminosity conversion given. The AGN component of their IR LF is also shown. As expected, their SED fitting shows that the

bolometric IR luminosity for AGN sources has a larger contribution from shorter wavelengths than for star-forming sources: they calculate $\log(L_{14 \mu\text{m}}/L_{\text{IR}}) = -1.19 \pm 0.10$ for star-forming sources, but just -0.57 ± 0.19 for AGN-dominated systems.

Being IR-selected, Wu et al.’s sample has the advantage of having similar properties to our own, and therefore we expect that the AGN contribution, calculated as a function of IR luminosity, will be applicable to our data. The good agreement between the two IR LFs – our own and Wu et al.’s – confirms this, and suggests that the AGN fraction derived from their sample is indeed applicable to our own. As a qualitative check, we compare the AGN contribution derived here to a previous result, the AGN LF at $12 \mu\text{m}$ as calculated by Rush, Malkan & Spinoglio (1993).

As discussed above, the AGN contribution to the star formation distribution only becomes significant at the bright end ($> 10^{12} L_{\odot}$), with the contribution from AGN being < 10 per cent for non-LIRGs, with SFRs below $\sim 15 M_{\odot} \text{ yr}^{-1}$. Both measures of AGN activity agree well, with a small (a factor of ~ 2) discrepancy at the low ($< 10^{11} L_{\odot}$) end – the low value of $\Phi(\text{AGN})$ at these luminosities renders this difference relatively unimportant. We do note, however, that the 5MUSES sample has spectral AGN estimators which are more accurate than the photometric estimators of the older work, and it is to this that we attribute the difference.

All LFs, star formation distributions and related values calculated hereafter have had the above AGN LF contribution subtracted as per the estimate from Wu et al., and are therefore interpretable in terms of pure star formation.

3.2 Luminosity functions

All galaxies in our data sets have been observed with multiple bands in the IR: either with *Spitzer* (at $24 \mu\text{m}/70 \mu\text{m}/160 \mu\text{m}$) or *IRAS* ($12 \mu\text{m}/25 \mu\text{m}/60 \mu\text{m}/100 \mu\text{m}$). As such, very accurate estimates – to better than 1 per cent for most galaxies – of their bolometric IR flux (defined as the integrated flux from 8 to $1000 \mu\text{m}$) can be made, using the three-component prescriptions provided by Dale & Helou (2002).

Fig. 2 shows the total infrared (TIR) LF for both the large IR sample (red triangles) and the LVL data (green squares). As we

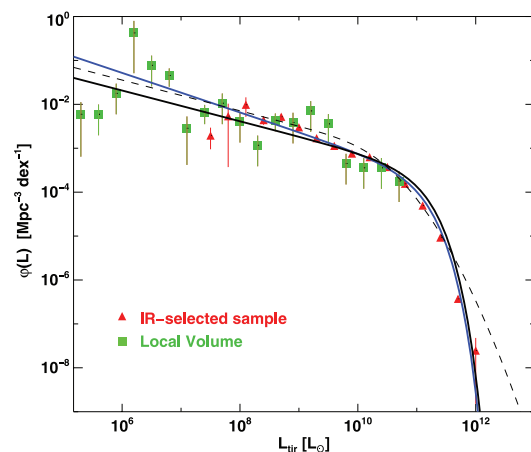


Figure 2. The bolometric IR LF for the IR sample. $1/V_{\text{max}}$ -derived points are shown for the IR-selected sample (red triangles) and the Local Volume galaxies (green squares). The black line is a least-squares Schechter fit to these points and the blue line is the maximum likelihood Schechter function fit. Errors (1σ) were calculated using the Monte Carlo bootstrapping. The dashed black line is the ‘canonical’ IRAS LF, from Takeuchi et al. (2003).

are interested in measuring the true underlying distributions, rather than just the local density enhancement, the number densities for the LVL data in Fig. 2 (and for all LVL volume densities hereafter) have been adjusted downwards by a factor of 1.85 as discussed above. The data points were derived using the $1/V_{\max}$ method, and the black line is the best-fitting Schechter function, obtained using standard least-squares fitting. The blue fit to the data is the Schechter function obtained using maximum likelihood fitting, as discussed above. The two methods agree on all counts, with the exception of a small discrepancy in the faint-end slope (1.41 ± 0.09 for the $1/V_{\max}$ method, 1.53 ± 0.08 for maximum likelihood).

For comparison, the LF for *IRAS* IR-selected galaxies from Takeuchi et al. (2003) (corrected slightly for our cosmological parameters) has been overplotted as a dashed line. There is a discrepancy at the high end, caused by both AGN removal (cf. our original sample shown in Fig. 1) and the redshift cut (taking $z < 0.1$ reduces the number of rare bright objects). Indeed, Takeuchi et al. attribute the ‘bright end bump’ seen in both our samples to the increasingly dominant AGN contribution at luminosities $L_{\text{IR}} > 5.0 \times 10^{11} L_{\odot}$ (the point at which their power-law LF diverges from our Schechter LF), which we have removed as per Section 3.1. Apart from this difference, the LF of the parent sample is essentially identical to our UV-matched subsample, suggesting that we have not introduced undue bias by insisting on a UV observation for inclusion in the sample (see Section 3.2.1).

Fig. 3 shows the FUV LF for the large UV sample (blue triangles) and the LVL data (green squares). The fitted black and blue curves are again, respectively, the least-squares fit to the $1/V_{\max}$ data points, and the maximum likelihood fit. Being the sample from Buat et al. (2007), the UV LF is identical to theirs. The dashed line is the field FUV LF from *GALEX*, presented by Wyder et al. (2005) (again corrected for the more recent cosmology). Again, the close similarity between our sample and the canonical ‘field’ LF suggested that our data do provide a representative sample of star-forming galaxies in the local Universe. We note, however, that there is a discrepancy with the LF of Wyder et al. (2005) at the UV-bright end. Again, the two LF derivation methods produce similar fits – the faint-end slope is 1.23 ± 0.09 for the $1/V_{\max}$ method and 1.31 ± 0.09 for maximum likelihood.

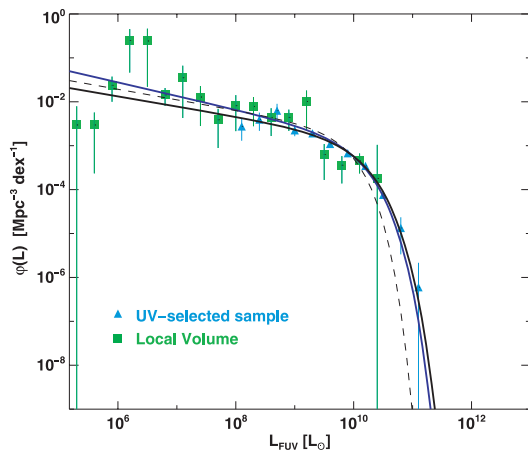


Figure 3. The FUV LF for the UV sample. $1/V_{\max}$ -derived points are shown for the UV-selected sample (blue triangles) and the Local Volume galaxies (green squares). The black line is a least-squares Schechter fit to these points and the blue line is the maximum likelihood Schechter function fit. Errors (1σ) were calculated using the Monte Carlo bootstrapping. The black dashed line is the UV LF from Wyder et al. (2005).

3.2.1 Non-detections, upper limits and incompleteness considerations

The construction of all the LFs and star formation distribution functions was carried out both by removing all galaxies with only upper flux limits (UV flux limits in the IR-selected sample, and vice versa) and by treating the upper flux limit as a detection equal to the limiting flux magnitude – as in Martin et al. (2005), the results in all cases do not depend on whether the galaxies are included, so we opted to include them.

This insensitivity to the inclusion of non-detections does not extend to the distribution of obscured star formation, as traced by the ratio $L_{\text{IR}}/L_{\text{FUV}}$. This parameter is highly sensitive to low flux values, such as those assumed from an upper limit (being a ratio, low UV fluxes lead to high values of $L_{\text{IR}}/L_{\text{FUV}}$), and is discussed more in Section 3.5.

We also have to consider the effect of sample incompleteness on the form of our SFR function. The large IR and UV data sets will suffer incompleteness starting at $\sim 0.1 M_{\odot} \text{ yr}^{-1}$ (note the downturns the IR and UV samples shown in Figs 2 and 3 starting at $\sim 10^8 L_{\odot}$). This is, however, the SFR at which the LVL data are highly complete, and hence our resultant function will remain statistically robust here.

At the faintest end of our function, however, we run into LVL incompleteness which cannot be ignored. Lee et al. (2009a) calculate the LVL data set to be complete to $\log \text{SFR} = -2.5 M_{\odot} \text{ yr}^{-1}$ based on the resultant $\text{H}\alpha$ flux limits of the sample. Below this level, too, small number statistics lead to derived SFRs being somewhat uncertain. We do not truncate our sample below this SFR, but it must be noted that values of Φ calculated at the faint end are inherently uncertain, and for reference we mark this regime on our star formation function shown in Fig. 4. We note, however, that calculations of the faint-end slope do not depend on this uncertain regime, and all derived parameters do not change if we fit only to the statistically complete data.

It is worth noting that the volumes probed by our field IR and UV data sets are large enough to avoid bias due to the large-scale structure. We can define a ‘correlation volume’, derived from the

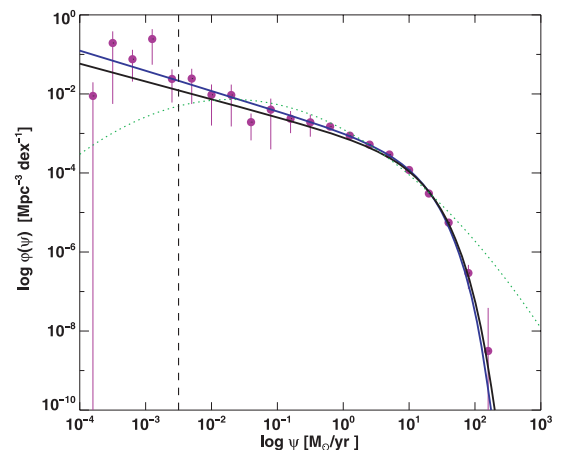


Figure 4. The SFR distribution function for the resultant combined sample, as described in Section 3.3. The black line is a least-squares Schechter function fit to these points and the blue line is the maximum likelihood Schechter function fit. The vertical dashed line is drawn at $\log \text{SFR} = -2.5 M_{\odot} \text{ yr}^{-1}$, the level at which incompleteness becomes significant. The green lognormal function indicates the SFR function given by Martin et al. (2005). Errors (1σ) were calculated using the Monte Carlo bootstrapping.

correlation length, which is an estimate of the absolute minimum volume you need to probe in order to be reasonably free of bias due to structure. We estimate this to be 500 Mpc^3 (approximating the correlation radius as 5 Mpc ; Zehavi et al. 2005). For a flux-limited survey, the volume probed will vary as a function of luminosity; at a luminosity of $10^8 L_\odot$ – the lower limit of our field data – our IR and UV data sets, respectively, probe volumes of 20 and 80 times this minimum ‘correlation volume’.

3.3 The distribution of star formation

For each galaxy in our sample, we derived a SFR as follows. The IRX, defined as

$$\text{IRX} = \log \left(\frac{L(\text{TIR})}{L(\text{FUV})_{\text{obs}}} \right),$$

is used as a measure of the internal dust absorption. This then can be used to accurately estimate the UV attenuation; it is known that the value of IRX is a good tracer of UV attenuation, and – importantly – remains robust independent of the details behind the extinction, such as geometry and dust properties (see Meurer, Heckman & Calzetti 1999). Conversions between IRX and $A(\text{FUV})$ in the literature are given by Calzetti et al. (2000), Kong et al. (2004), Buat et al. (2005) and Cortese, Gavazzi & Boselli (2008). For the purposes of this paper, we use the IRX– $A(\text{FUV})$ relation derived by Burgarella, Buat & Iglesias-Páramo (2005):

$$A(\text{FUV}) = -0.028x^3 + 0.392x^2 + 1.094x + 0.546,$$

where x is the IRX defined above.

Once the observed UV luminosity has been corrected for dust attenuation using the above expression, we converted to a SFR as per the method described by Iglesias-Páramo et al. (2006):

$$\log \text{SFR} [\text{M}_\odot \text{ yr}^{-1}] = \log(L_{\text{FUV,corr}} [L_\odot]) - 9.51.$$

This calibration uses a Salpeter IMF from 0.1 to 100 M_\odot , was derived specifically for the *GALEX* bands and is similar to other estimators in the literature. For comparison, the Kennicutt (1998) conversion factor:

$$\text{SFR} [\text{M}_\odot \text{ yr}^{-1}] = 1.4 \times 10^{-28} L_v [\text{erg s}^{-1} \text{ Hz}^{-1}]$$

when converted from a monochromatic UV luminosity – taking the effective central wavelength of the *GALEX* FUV filter to be 1532 \AA – gives $\log \text{SFR} [\text{M}_\odot \text{ yr}^{-1}] = \log(L_{\text{FUV,corr}} [L_\odot]) - 9.56$: ~ 12 per cent lower than our value. We adopt the former conversion, as it was derived specifically for the *GALEX* filters by Iglesias-Páramo et al. (2006). We note that this method does carry some implicit assumptions, such as a constant star formation history, and a lack of stochasticity in populating the upper (UV-emitting) regions of the IMF. As noted above, for the purpose of calculating SFRs we include non-detections, by assigning a flux equal to the detection limit – the overall results do not change if we instead discard all non-detected galaxies.

We then generated a ‘resultant’ sample from our large UV and IR data sets using the ‘incoherent combination of domain-independent samples’ method (as described by Avni & Bahcall 1980). Using this method, both samples can be considered together: the combined sample is split into two ‘domains’, defined by 1: $[S_{\text{IR}} > S_{\text{IR}}^{\text{lim}}]$ and $S_{\text{UV}} > S_{\text{UV}}^{\text{lim}}$ and 2: $[S_{\text{IR}} < S_{\text{IR}}^{\text{lim}}]$ and $S_{\text{UV}} > S_{\text{UV}}^{\text{lim}}$. Any duplicates must be dealt with, as each galaxy is clearly just a single probe of the population. The duplicate is removed, and the remaining galaxy is assigned whichever value of V_{max} (i.e. based on either the IR or UV limit) is *larger* – this is because for the object to be included, it is

sufficient that it satisfies the weaker selection criteria. Objects in the first domain then have V_{max} values calculated as normal. Objects in the second domain, being constrained by having an IR flux below the limit of the IR-selected sample (and thus only appearing in the UV-selected sample), have a reduced region of parameter space to exist in, and have their values of V_{max} defined by $(V_{\text{max,IR}} - V_{\text{max,UV}})$.

We cannot apply the same technique to combine with the Local Volume data, as the LVL sample is designed to be volume limited and therefore has a less cleanly defined selection function; where the samples overlap with LVL, we simply weight the two data sets inversely by their bootstrap-derived errors. The final, combined sample consists of 10 704 galaxies (257 from the LVL, 562 from the UV and 10 141 from the IR, where duplicate galaxies have been removed as explained above).

The SFR function is shown in Fig. 4. The purple data points are V_{max} -derived points for the combined sample as described above. We use least-squares fitting to fit a standard Schechter function to the points: the best-fitting parameters are found to be $(\Phi^* = 0.00015 \pm 0.0003 \text{ Mpc}^{-3}, \psi^* = 9.0 \pm 0.3 \text{ M}_\odot \text{ yr}^{-1}, \alpha = -1.48 \pm 0.07)$. This Schechter function is shown in Fig. 4 as a black line.

We also used the maximum likelihood method described above to find the best-fitting Schechter function to the combined sample; the best-fitting parameters are found to be $(\Phi^* = 0.00016 \pm 0.0004 \text{ Mpc}^{-3}, \psi^* = 9.2 \pm 0.3 \text{ M}_\odot \text{ yr}^{-1}, \alpha = -1.51 \pm 0.08)$, and this is plotted in Fig. 4 as a blue line. The two methods of calculating the SFR function match closely, suggesting that the V_{max} points are not overly biased by the presence of clustering. Hereafter, we adopt the parameters of the maximum likelihood derived function.

The faint-end slope of -1.51 continues monotonically almost to the limits of the data, until the low galaxy numbers available at SFRs $< 10^{-3} \text{ M}_\odot \text{ yr}^{-1}$ lead to the degradation of the relation due to noise. This is in contrast to previous extrapolations (e.g. Martin et al. 2005) which predicted a lognormal form for the SFR distribution function, with a maximum value of $\Phi(\psi)$ at $\psi \sim 10^{-2} \text{ M}_\odot \text{ yr}^{-1}$ and a gradual decline thereafter.

This result is particularly interesting in the light of theoretical predictions of structure formation. There has been a long-running conflict between the small-scale predictions for DM haloes, and the observational results of the galaxies that dwell inside them (Klypin et al. 1999). Galaxy formation models predict a scale-invariant form for DM, which clusters hierarchically even on the smallest scales; this manifests as a steep faint-end slope to the mass function, of $\alpha \sim -1.8$. Observations of galaxy LFs, however, have found a dearth of small-scale objects at the bottom of the LF, indicating a much shallower faint-end slope. Either the predictions for the DM mass function are incorrect, or the complex non-linear processes involved in galaxy formation conspire to suppress the formation of baryonic structures on the smallest scales.

One suggestion has been that the lack of satellite-scale galaxies results from star formation being systematically suppressed on the smallest scales (i.e. Bullock, Kravtsov & Weinberg 2000; Robertson et al. 2005); the form of our SFR distribution function, however, is consistent with the idea that this is not the case: a monotonically increasing faint-end slope to the SFR distribution function suggests that any process suppressing star formation must be operating in a scale-free manner with respect to halo mass.

3.4 The star formation rate volume density

Using the distribution of SFR, it is possible to calculate the distribution of SFR *volume density*. We define the distribution function

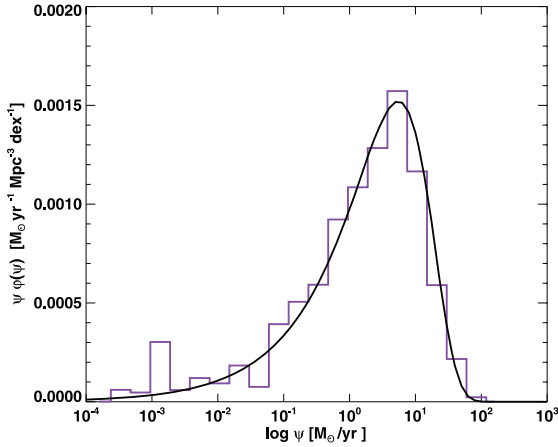


Figure 5. The SFR volume density distribution function, for the resultant combined sample as in Fig. 4. The purple histogram shows the V_{max} -derived data as above, and the black fit to the data is the convolved Schechter function $\psi \Phi(\psi)$, with the maximum likelihood fit parameters as above.

of SFR density as

$$\psi \Phi(\psi) = \psi \Phi^* \left(\frac{\psi}{\psi^*} \right)^\alpha e^{-(\psi/\psi^*)} d \left(\frac{\psi}{\psi^*} \right).$$

Fig. 5 shows the distribution of SFR volume density, for the ‘combined’ sample shown in Fig. 4, along with the convolved Schechter fit for the parameters given above.

We can therefore integrate $\psi \Phi(\psi)$ to give the total SFR volume density in the local Universe,

$$\rho_{\text{SFR}} = \int_0^\infty d\psi \psi \Phi(\psi),$$

which for the parameters of the Schechter fit to our data ($\Phi^* = 0.00016 \pm 0.0004 \text{ Mpc}^{-3}$, $\psi^* = 9.2 \pm 0.3 \text{ M}_\odot \text{ yr}^{-1}$, $\alpha = -1.51 \pm 0.08$) gives $\rho_{\text{SFR}} = (25 \pm 1.7) \times 10^{-3} \text{ M}_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}$.

This is in good agreement with most recent derivations of this result (see Table 1 for a compilation of recent results). There is a relatively large spread in the derived values of the SFR volume density – greater than a factor of 2, beyond the errors quoted on the individual measurements. This is discussed briefly by Serjeant et al. (2002) [who derive their own 1.4 GHz based value of $(21 \pm 5) \times 10^{-3} \text{ M}_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}$], who attribute the discrepancy to a systematic underestimation of the extinction using the Balmer decrement in some emission-line-based studies.

The total value of ρ_{SFR} can also be decomposed into ‘UV’ and ‘IR’ components, by integrating the value of $\psi \Phi(\psi)$ derived from each component individually. Doing so leads to values of $\rho_{\text{SFR}}(\text{IR}) = 0.011 \text{ M}_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}$, and $\rho_{\text{SFR}}(\text{UV}) = 0.012 \text{ M}_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}$. The LVL contribution is $0.0007 \text{ M}_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}$. This is 47 per cent, 50 per cent and 3 per cent of the total for the IR, UV and LVL components, respectively. This result – that about half of the energy from the total cosmic star formation budget is re-processed by dust – is well known, and is in line with previous studies. Takeuchi, Buat & Burgarella (2005) found that of their derived total SFR volume density ($19 \times 10^{-3} \text{ M}_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}$), 56 per cent was from dust-reprocessed emission.¹ For consistency (and because our statistical AGN removal involves some uncertainty), we have checked

the value of ρ_{SFR} calculated from the sample *without* the statistical correction for AGN contamination (as per Section 3.1). As the correction is only significant at the upper end (beyond ψ^*), the value only changes slightly: without any AGN correction applied, we calculate $\rho_{\text{SFR}} = (26 \pm 2.2) \times 10^{-3} \text{ M}_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}$.

We may also compute the fraction of the local cosmic SFR density occurring in starburst environments. For the purposes of such an analysis, we define a starburst as a system forming stars at $\geq 10 \text{ M}_\odot \text{ yr}^{-1}$. Using the SFR density distribution, we can thus integrate from $10 \text{ M}_\odot \text{ yr}^{-1}$ to infinity:

$$f_{\text{burst}} = \frac{1}{\rho_{\text{SFR}}} \int_{\psi_{\text{burst}}}^\infty d\psi \psi \Phi(\psi).$$

For our data, this value is $0.0049 \pm 0.00039 \text{ M}_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}$, or 20.4 per cent of the total SFR volume density; by our (admittedly somewhat crude) definition, one fifth of the star formation in the local Universe is provided by starbursts. This is consistent with the values found by Brinchmann et al. (2004) using specific SFRs from Sloan Digital Sky Survey (SDSS). Interestingly, Lee et al. (2009a) also find that 20 per cent of star formation in the dwarf galaxy population is concentrated in high $H\alpha$ equivalent width systems.

There are many interesting values that can be derived from the distribution shown in Fig. 5, including the starburst fraction (discussed above), the ‘dividing’ SFR at which 50 per cent of the star formation is happening both above and below, and so on. Rather than providing a list of values for various integration limits, it is more enlightening to consider the behaviour of the cumulative fraction of SFR volume density, which is shown in Fig. 6. This shows SFR, plotted against the fraction of the total star formation volume density coming from SFRs *lower* than that SFR. The data show a power-law increase in SFR volume density fraction, over 5 orders of magnitude until the truncation at $\sim 20 \text{ M}_\odot \text{ yr}^{-1}$. From this it can be seen that the ‘50 per cent’ divide occurs at $\sim 3 \text{ M}_\odot \text{ yr}^{-1}$, about the SFR of the Milky Way (e.g. Misiriotis et al. 2006).

It is also interesting to consider the contribution to the total SFR volume density from LIRGs and ULIRGs. These IR-bright galaxies (defined as having $L_{\text{IR}} > 10^{11} L_\odot$ and $> 10^{12} L_\odot$, respectively) are rare in the local Universe, but become more and more important with lookback time, becoming an increasingly dominant component of the total SFR volume density at higher redshifts (Schiminovich et al. 2005; Magnelli et al. 2009; Goto et al. 2010).

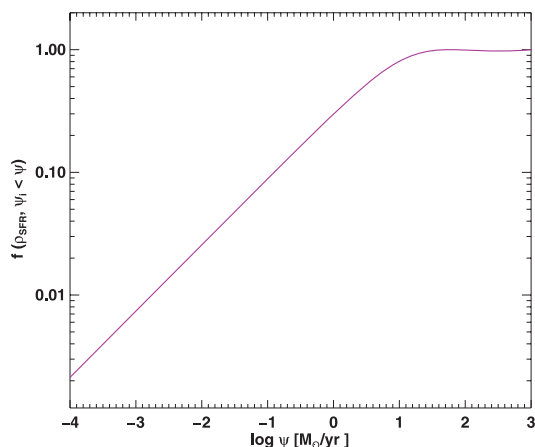
Fig. 7 shows the SFR distribution function shown in Fig. 5, expanded along a second dimension with TIR luminosity [the colour-coded ‘z’ axis corresponds to the value of $\psi \Phi(\psi)$]. Horizontal lines have been drawn at the two characteristic luminosity cuts for LIRGs and ULIRGs, to illustrate the total contribution to the local SFRD coming from those galaxies. By integrating the star formation rate density (SFRD) function above and below the cut-off lines, we can estimate the contribution to the total from both LIRGs and ULIRGs; 9 ± 1 per cent of total star formation is occurring in LIRGs, while just 0.6 ± 0.2 per cent is occurring in ULIRGs. This is in good agreement with literature values – Goto et al. (2010), for example, reach similarly small estimates of 7 ± 1 per cent for LIRGs, and 0.4 ± 0.1 per cent for ULIRGs. As discussed above, the dominant object type driving the total star formation in the local Universe are normal, secularly evolving galaxies with SFRs comparable to the Milky Way – despite their prodigious SFRs, the sparsity of LIRGs/ULIRGs means that they do not contribute significantly.

Due to the (U)LIRGs’ bright IR luminosities, the derived LIRG/ULIRG fractions are highly sensitive to the nature of the AGN correction used, which strongly affects the behaviour of the LF at the bright end. If we do not correct for AGN contamination as

¹ See also Lagache, Puget & Dole (2005), and references therein.

Table 1. Derivations from the literature of the SFR volume density in the local Universe.

Reference	SFR tracer	$\langle z \rangle$	SFRD ($10^{-3} \text{ M}_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-1}$)
Gallego et al. (2002)	[O II]	0.025	9.3 ± 3
Sullivan et al. (2000)	[O II]	0.15	23 ± 3
Hogg et al. (1998)	[O II]	0.20	11 ± 4
Gallego et al. (1995)	H α	0.022	12 ± 5
Tresse & Maddox (1998)	H α	0.2	25 ± 4
Sullivan et al. (2000)	H α	0.15	14 ± 3
Pérez-González et al. (2003)	H α	0.025	25 ± 4
Ly et al. (2007)	H α	0.08	13 ± 4
Hanish et al. (2006)	H α	0.01	16^{+2}_{-4}
Brinchmann et al. (2004)	H α	0.15	29 ± 5
Dale et al. (2010)	H α	0.16	10^{+6}_{-4}
Westra et al. (2010)	H α	0.05	6 ± 2
Westra et al. (2010)	H α	0.15	12 ± 3
Serjeant, Gruppioni & Oliver (2002)	1.4 GHz	0.005	21 ± 5
Condon (1989)	1.4 GHz	0.005	21 ± 0.5
Sullivan et al. (2000)	FUV	0.150	39 ± 5
Martin et al. (2005)	FUV+IR	0.02	21 ± 2
This work	FUV+IR	0.05	25 ± 1.6

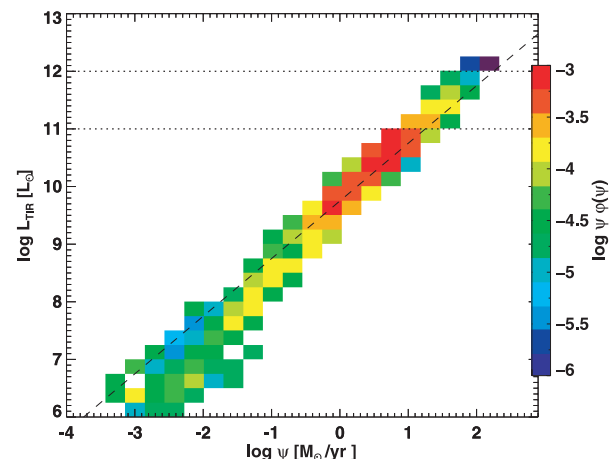

Figure 6. The cumulative fraction of SFR volume density.

per Section 3.1, the fractional contribution to the total SFRD from LIRGs and ULIRGs, respectively, is 14 ± 2 per cent and 1.5 ± 0.4 per cent. It should be noted, then, that our original derived (U)LIRG contributions are highly dependent on the AGN correction.

3.5 Dust-obscured star formation

It is possible to examine the distribution function of extinction, in much the same way as we have previously examined the distribution function of luminosity and star formation. This will examine the global behaviour of dust-obscured star formation, as a function of the amount of obscuration. We use as our measure of ‘extinction’ or ‘dust obscuration’ the ratio of IR to observed UV luminosities, IRX, as described in Section 3.3. Having a strong positive correlation with both luminosity and SFR (Wang & Heckman 1996; Hopkins et al. 2001; Buat et al. 2005; Takeuchi et al. 2010) the IRX distribution function should resemble the Schechter-like distribution functions derived elsewhere in this work.

Fig. 8 shows the distribution functions of IRX for the three samples included in this work. The three samples have been shown


Figure 7. The SFR volume density function $\psi \Phi(\psi)$ (as in Fig. 5) further expanded in a second dimension (along the ordinate) to show the breakdown with TIR luminosity. The colour-coded ‘ z ’ axis represents $\psi \Phi(\psi)$. Horizontal lines are drawn at $\log(L) = \{11, 12\}$, the respective defined minimum luminosities for LIRGs and ULIRGs. The dotted line shows the $L(\text{FIR})$ –SFR scaling given by Iglesias-Páramo et al. (2006).

separately, as a ‘resultant’ sample, constructed by combining the different samples with different selection functions, has little physical interpretation in this case. The distribution of IRX, and the difference between the samples, sheds light on FIR and FUV selection effects (Xu et al. 2006) as much as it does the underlying physical properties of the galaxies in question.

It is important to note here that, in contrast to the work with luminosity and SFR described above, the choice to include/exclude non-detections significantly alters the shape of the IRX distribution function, simply because the IRX is defined as a ratio of luminosities. So, while a non-detection in the UV has little effect in terms of the SFR function (it simply implies that the bolometric output is predominantly in the IR, and the UV is negligible), it has the result of making the value of IRX formally infinite. If non-detections are assumed to have flux equal to the limiting flux (as above), the

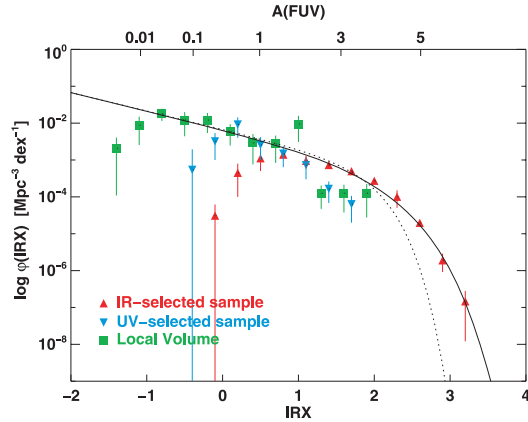


Figure 8. The distribution function of IRX for the three samples. The red, blue and green lines are, respectively, the IR-selected, UV-selected and LVL data sets. The solid black line is the best-fitting modified Schechter fit (see text), and the dotted black line shows the same parameters converted into a standard Schechter fit by setting the high-end slope $\beta = 1$. It can be seen that a normal Schechter function is a poor fit to the data. The equivalent FUV attenuation has been shown on the upper x-axis, calculated as detailed in the text.

IRX distribution function becomes essentially meaningless, as it is driven entirely by the extreme upper/lower limit values of IRX assigned semi-arbitrarily to the non-detections. We therefore chose to *exclude* non-detections for the purpose of this IRX analysis. The number of galaxies thus excluded from the IR and UV samples is, respectively, 1784 and 73. As a result of excluding the non-detections, the distribution functions shown in Fig. 8 can be interpreted as lower bounds on the IRX distribution functions.

All three samples exhibit broadly similar behaviour in their value of $\Phi(\text{IRX})$: a decrease in Φ towards higher values of IRX [analogous to the decrease in $\Phi(L)$ for rare high-luminosity systems], and a second dip at the lowest values of IRX probed. With the exception of an overdensity at $\text{IRX} = 1$ in the LVL sample, the three samples do share a common Schechter-like ‘envelope’, though the individual samples do differ. The UV and LVL samples fall off steeply between an IRX of 1 and 2. In contrast, the high-IRX end (entirely comprising the IR-selected sample) is much shallower than would be expected from a standard Schechter function, however, and is best fit with a modified Schechter function, which has been adapted following Hopkins et al. (2010) to leave the high-IRX end slope β as a free parameter:

$$\Phi(X)dX = \Phi^* \left(\frac{X}{X^*} \right)^\alpha \exp[-(X/X^*)^\beta] d\left(\frac{X}{X^*} \right),$$

where a high-end slope $\beta < 1$ allows for a shallower fall-off at high IRX than a standard Schechter function ($\beta = 1$).

This is plotted in Fig. 8 as a solid black line (the standard Schechter fit using the same parameters is, for comparison, shown as a dotted line). This has a low-end slope of $\alpha = -1.51 \pm 0.11$, and a high-end slope $\beta = 0.63 \pm 0.09$. This is significantly shallower than 1, which would be expected given the previously noted strong correlation between luminosity/SFR and IRX, and deserves further examination. We used NED to look up the high-IRX galaxies in our sample, which make up the top of the distribution. The galaxies inhabiting this extended shallow ‘tail’ to the distribution are, predominantly, high-inclination discs (see Fig. 9 for a selection), which have higher column densities of dust than face-on galaxies.

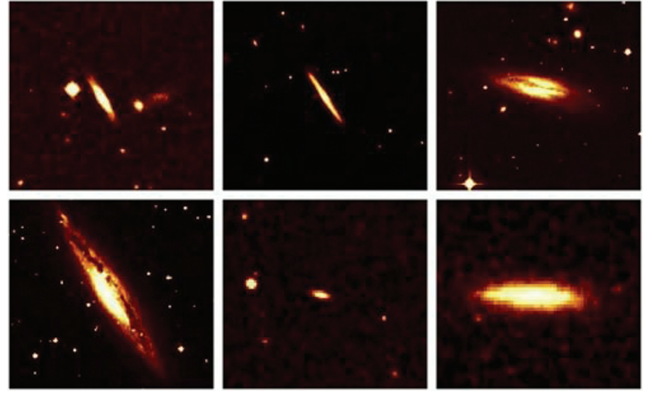


Figure 9. A selection of the high-IRX galaxies from the IR-selected sample, showing their unusual/edge-on morphology and prominent dust lanes. Clockwise from the top left: IRAS F14299+3631; F01221+0944; F21596-1909; F14263+2555; F13228+1837 and F11290-3001. Images obtained from NED.

The shallow high-end slope seems to be driven by inclination (and its associated radiative transfer effects) as much as inherent, angle-averaged extinction, which is one possible reason for the disparity between the LF and the IRX distribution function. The deviation from a Schechter function is, to some extent, expected. The Press–Schechter formalism describes the distribution of DM halo masses – parameters that are well correlated with halo mass should have a Schechter-like scale distribution. As the parameter in question becomes further removed from the underlying halo mass, there is less reason to expect that a Schechter function will be a good model of the parameter distribution. The additional complication of complex, inclination-dependent radiative transfer effects causes the IRX distribution to take a non-Schechter form.

It is again possible to break down the SFR volume density distribution (shown in Fig. 5) with IRX, to see at which extinction level the contribution is the greatest. This is shown in Fig. 10. The SFR volume density function is decomposed into bins of IRX along the y-axis, while the colour-coded ‘z’ axis corresponds to $\psi \Phi(\psi)$. The

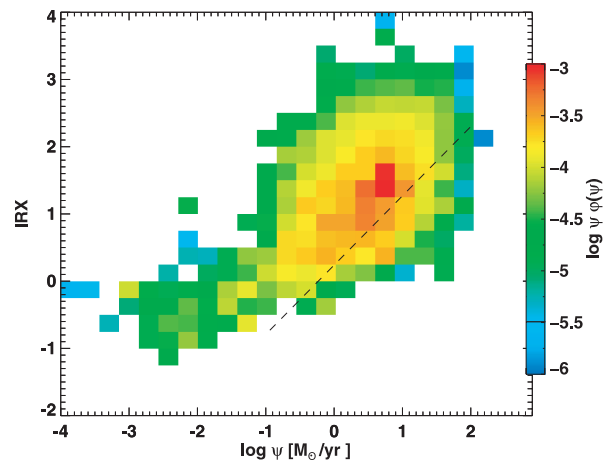


Figure 10. The SFR volume density function as a function of SFR (Fig. 5), further expanded in a second dimension (along the ordinate) to show the breakdown with IRX. The ‘z’ axis represents $\psi \Phi(\psi)$. The dotted line shows the IRX–SFR relationship derived for star-forming $z = 0$ galaxies given by Overzier et al. (2011).

peak in the SFRD function which occurs at $\psi \sim 3 \text{ M}_{\odot} \text{ yr}^{-1}$ occurs at $\text{IRX} \sim 1.4$ in this breakdown. While the SFRD function decreases sharply towards higher SFRs, however, the decline towards higher values of IRX is much more shallow, with a non-trivial fraction of the total SFRD coming from highly extincted systems. 13 ± 1 per cent of the total SFR volume density comes from systems with $\text{IRX} > 2$, dusty systems with a factor of 100 difference between their IR and UV luminosities.

4 CONCLUSIONS

We have presented an analysis of the IR- and UV-selected LFs for galaxies in the $z \sim 0$ Universe. We make use of large survey data sets ($>10\,000$ galaxies) from the *IRAS* and *GALEX* satellites to provide us with good number statistics; we also utilize the LVL survey to extend the field relations 1–2 orders of magnitude deeper than before, probing the faintest visible star-forming galaxies. Our main conclusions are as follows:

- (i) The distribution function of star formation in the local Universe has a faint-end slope of $\alpha = -1.51 \pm 0.07$. This appears to be constant and monotonic as far back as the data allow us to go, down to the faintest dwarf galaxies with SFRs $< 0.001 \text{ M}_{\odot} \text{ yr}^{-1}$.
- (ii) The SFR volume density of the local ($z \sim 0$) Universe is found to be $0.025 \pm 0.0016 \text{ M}_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}$. The distribution function of SFR volume density shows that this value is predominantly driven by modest star-forming galaxies, with SFRs of $\sim 3 \text{ M}_{\odot} \text{ yr}^{-1}$, similar to the Milky Way. ‘Starburst’ galaxies, with SFRs $\geq 10 \text{ M}_{\odot} \text{ yr}^{-1}$, contribute 20 per cent to the local cosmic SFR density.
- (iii) In contrast to the high- z Universe, LIRGs and ULIRGs play only a modest role in determining the total SFR volume density, contributing 9 ± 1 and 0.7 ± 0.2 per cent, respectively.
- (iv) The distribution function of ratio $L_{\text{IR}}/L_{\text{FUV}}$, ‘ IRX ’, behaves somewhat differently to that of star formation, despite the strong correlation. The low- IRX slope is comparable, but the high- IRX slope is significantly shallower, and is very poorly fit by a standard Schechter function. This is partly attributable to the addition of geometric effects (whereby edge-on galaxies have inflated extinctions relative to their face-on counterparts), and varying the bright end slope β to the data provides a good fit for $\beta = 0.63 \pm 0.09$.
- (v) Breaking down $\psi \Phi(\psi)$ with IRX , it can be seen that the peak of SFRD occurs at $\text{IRX} \sim 1.4$. Extremely dusty systems contribute a minor but significant fraction to the total SFRD, with 13 ± 1 per cent coming from dusty galaxies with $\text{IRX} > 2$.

ACKNOWLEDGMENTS

We would like to thank Matt Malkan for helpful and enlightening discussions. This research has made use of the NED which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. We acknowledge the usage of the HyperLeda data base (<http://leda.univ-lyon1.fr>). MB is supported by STFC.

REFERENCES

Avni Y., Bahcall J. N., 1980, *ApJ*, 235, 694
 Barkana R., 2002, *New Astron.*, 7, 337
 Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, *MNRAS*, 398, 1150
 Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, *MNRAS*, 351, 1151

Buat V. et al., 2005, *ApJ*, 619, L51
 Buat V. et al., 2007, *ApJS*, 173, 404
 Bullock J. S., Kravtsov A. V., Weinberg D. H., 2000, *ApJ*, 539, 517
 Burgarella D., Buat V., Iglesias-Páramo J., 2005, *MNRAS*, 360, 1413
 Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *ApJ*, 533, 682
 Calzetti D. et al., 2010, *ApJ*, 714, 1256
 Condon J. J., 1989, *ApJ*, 338, 13
 Cortese L., Gavazzi G., Boselli A., 2008, *MNRAS*, 390, 1282
 Dale D. A., Helou G., 2002, *ApJ*, 576, 159
 Dale D. A. et al., 2009, *ApJ*, 703, 517
 Dale D. A. et al., 2010, *ApJ*, 712, L189
 Efsthathiou G., Ellis R. S., Peterson B. A., 1988, *MNRAS*, 232, 431
 Fu H. et al., 2010, *ApJ*, 722, 653
 Gallego J., Zamorano J., Aragon-Salamanca A., Rego M., 1995, *ApJ*, 455, L1
 Gallego J., García-Dabó C. E., Zamorano J., Aragón-Salamanca A., Rego M., 2002, *ApJ*, 570, L1
 Gil de Paz A. et al., 2007, *ApJS*, 173, 185
 Goto T. et al., 2010, *A&A*, 514, 6
 Hanish D. J. et al., 2006, *ApJ*, 649, 150
 Hogg D. W., Cohen J. G., Blandford R., Pahre M. A., 1998, *ApJ*, 504, 622
 Hopkins A. M., 2004, *ApJ*, 615, 209
 Hopkins A. M., Connolly A. J., Haarsma D. B., Cram L. E., 2001, *AJ*, 122, 288
 Hopkins P. F., Younger J. D., Hayward C. C., Narayanan D., Hernquist L., 2010, *MNRAS*, 402, 1693
 Iglesias-Páramo J. et al., 2006, *ApJS*, 164, 38
 Karachentsev I. D., Makarov D. A., 1996, *AJ*, 111, 794
 Karachentsev I. D., Karachentseva V. E., Huchtmeier W. K., Makarov D. I., 2004, *AJ*, 127, 2031
 Kennicutt R. C., Jr, 1998, *ARA&A*, 36, 189
 Kennicutt R. C., Jr, Lee J. C., Funes José G. S. J., Sakai S., Akiyama S., 2008, *ApJS*, 178, 247
 Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, *ApJ*, 522, 82
 Kong X., Charlot S., Brinchmann J., Fall S. M., 2004, *MNRAS*, 349, 769
 Lagache G., Puget J., Dole H., 2005, *ARA&A*, 43, 727
 Lee J. C., Kennicutt R. C., José G., Funes S. J., Sakai S., Akiyama S., 2009a, *ApJ*, 692, 1305
 Lee J. C. et al., 2009b, *ApJ*, 706, 599
 Lee J. C. et al., 2011, *ApJS*, 192, 6
 Liske J., Lemon D. J., Driver S. P., Cross N. J. G., Couch W. J., 2003, *MNRAS*, 344, 307
 Ly C. et al., 2007, *ApJ*, 657, 738
 Magnelli B., Elbaz D., Chary R. R., Dickinson M., Le Borgne D., Frayer D. T., Willmer C. N. A., 2009, *A&A*, 496, 57
 Marinoni C., Monaco P., Giuricin G., Costantini B., 1999, *ApJ*, 521, 50
 Martin D. C. et al., 2005, *ApJ*, 619, L59
 Martin D. C. et al., 2007, *ApJS*, 173, 415
 Meurer G. R., Heckman T. M., Calzetti D., 1999, *ApJ*, 521, 64
 Misiriotis A., Xilouris E. M., Papamastorakis J., Boumis P., Goudis C. D., 2006, *A&A*, 459, 113
 Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, *ApJ*, 524, L19
 Overzier R. A. et al., 2011, *ApJ*, 726, 7
 Pérez-González P. G., Gallego J., Zamorano J., Alonso-Herrero A., Gil de Paz A., Aragón-Salamanca A., 2003, *ApJ*, 587, L27
 Press W. H., Schechter P., 1974, *ApJ*, 187, 425
 Robertson B., Bullock J. S., Font A. S., Johnston K. V., Hernquist L., 2005, *ApJ*, 632, 872
 Rush B., Malkan M. A., Spinoglio L., 1993, *ApJS*, 89, 1
 Salim S., Rich R. M., Charlot S., Brinchmann J., Johnson B. D., Schiminovich D., 2007, *ApJS*, 173, 267
 Schechter P., 1976, *ApJ*, 203, 297
 Schiminovich D. et al., 2005, *ApJ*, 619, L47
 Serjeant S., Gruppioni C., Oliver S., 2002, *MNRAS*, 330, 621
 Sullivan M., Treyer M. A., Ellis R. S., Bridges T. J., Milliard B., Donas J., 2000, *MNRAS*, 312, 442

- Takeuchi T. T., Yoshikawa K., Ishii T. T., 2003, *ApJ*, 587, L89
Takeuchi T. T., Buat V., Burgarella D., 2005, *A&A*, 440, L17
Takeuchi T. T., Buat V., Heinis S., Giovannoli E., Yuan F., Iglesias-Páramo J., Murata K. L., Burgarella D., 2010, *A&A*, 514, A4
Tresse L., Maddox S. J., 1998, *ApJ*, 495, 691
Wang B., Heckman T. M., 1996, *ApJ*, 457, 645
Wang L., Rowan-Robinson M., 2009, *MNRAS*, 398, 109
Wang L., Rowan-Robinson M., 2010, *MNRAS*, 401, 35
Westra E., Geller M. J., Kurtz M. J., Fabricant D. G., Dell’Antonio I., 2010, *ApJ*, 708, 534
Wu Y. et al., 2011, *MNRAS*, accepted (arXiv:1104.0687)
Wyder T. K. et al., 2005, *ApJ*, 619, L15
Xu C. K. et al., 2006, *ApJ*, 646, 834
Yahil A., Strauss M. A., Davis M., Huchra J. P., 1991, *ApJ*, 372, 380
Yuan T., Kewley L. J., Sanders D. B., 2010, *ApJ*, 709, 884
Zehavi I. et al., 2005, *ApJ*, 630, 1

This paper has been typeset from a \LaTeX file prepared by the author.